## The gravity model reconsidered

van Maanen, T.

1985

# SERIE RESEARCHMEMORANDA

THE GRAVITY MODEL RECONSIDERED

T. van Maanen

Researchmemorandum 1985-18 Juni '85

VRIJE UNIVERSITEIT
EKONOMISCHE FAKULTEIT
AMSTERDAM

## The gravity model reconsidered

### Introduction

The gravity model consists of one basic equation. This equation is:

$$X_{ij} = f(GNP_i, GNP_j, D_{ij}) \tag{1}$$

in which $X_{ij}$ is a flow (of goods/ or persons) between places i and j. $GNP_{i(j)}$ stands for a push(pull) factor of i(j). The push indicates the number and amount of possible flows from i and the pull indicates the attraction of place j upon the flows originating from i. $D_{ij}$ indicates a resistance factor: eg. the distance between i and j. Sometimes more equations are added which represent restrictions upon the number of flows originating from i or arriving at j. The model is used in geographics as a spatial interaction model and in economics as a trade model. See Nijkamp(1975), Linnemann(1966) and Bennett(1985). The gravity model of international trade, in which $X_{ij}$ stands for trade in goods, GNP for gross national product and $D_{ij}$ for the distance between i and j, is seen by Linnemann as a reduced form equation. The supply and demand for internationally traded goods are set equal in an equilibrium equation. Later Alonso(1978) provided a more general derivation, which can be utilized for other movements. Traditionally the gravity model is estimated using OLS. In this paper it is shown that this technique can yield biased results. Therefore another estimator, the Tobit estimator, is utilized, which provided unbiased results. In literature, different procedures to compute the Tobit estimation are suggested. In this paper two different procedures are presented: those of Heckman and Fair(1977). Another element dwelt upon, is the explicit formulation of the balancing factor. In trade models prices are normally used as a balancing factor. In the reduced form prices are substituted by other variables. However, explicit formulation shows that an assumption about the price elasticities is needed in order to estimate the parameters of the equation when the equation does not include a price variable. This is shown by Linnemann and Alonso.

Review of literature

In this paragraph, some estimations are presented in which the gravity equation was used to measure properties in international trade. Linnemann's thesis can be regarded as a starting point for the use of the gravity equation in international trade. Referring to his research many authors used the gravity equation after the publication of his thesis. Some basic features of his equation are: (1) the use of a cross section procedure, (2) the use of a least squares estimator (mostly OLS), (3) the inclusion of gross national product (Y) and the population (N) of the trading countries and the distance (D) between the trading countries as independent variables, (4) the use of dummies to describe deviations from the normal pattern and (5) the transformation of the variables into natural logarithms.
Linnemann estimated merchandise trade ($X_{ij}$) as:

$$X_{ij} = 0.12 + \underset{(0.02)}{0.98} \log Y_i + \underset{(0.02)}{0.86} \log Y_j - \underset{(0.03)}{0.21} \log N_i - \underset{(0.03)}{0.14} \log N_j$$
$$- \underset{(0.03)}{0.77} \log D_{ij} + \underset{(0.14)}{1.27} \log P_{ij}^{UUC} + \underset{(0.26)}{2.57} \log P_{ij}^{FFC} + \underset{(0.67)}{6.89} \log P_{ij}^{B}$$

$P_{ij}^{UUC}$ = 2 if both countries (i,j) are members of the British Commonwealth and associates; $P_{ij}^{FFC}$ = 2 if both countries are members of the French Community and associates; $P_{ij}^{B}$ = 2 if the two countries belong to the group Zaire, Belgium or to the group Portugal, Mozambique and Angola. Data connected with zero trade flows were omitted from the estimation procedure.

Standard errors are written between parentheses.

The suffix i stands for the exporting country and j stands for the importing country. Linnemann regarded 80 exporting countries and 80 importing countries, so in principle 6400 trade flows could be estimated. The year for which estimations are made is 1959.

Hirch & Lev omitted the population variable (N) and they included a variable which indicated the difference in per capita income (pci). They wanted to test the Linder hypothesis which states that trade in manufactured goods is enlarged if both countries have an overlapping demand structure- for which the difference in pci is a proxy.

The dependent variable $X_{ijk}$ stands for food, textiles and clothing, machinery other than electric, chemicals, electrical equipment, electronics and miscellaneous manufactured articles. The number of exporting countries (i) was four, the number of importing (j) countries was 111. The year for which estimates are made is 1966. The results:

$$\log X_{ijk} = a + 1.74 \log Y_i + 0.76 \log Y_j - 0.74 \log D_{ij}$$

$$\underset{(0.06)}{-\ 0.448} \log \left( \min\left(\left(\frac{Y_i/N_i}{Y_j/N_j}\right), \left(\frac{Y_j/N_j}{Y_i/N_i}\right)\right)\right)$$

Fortune estimated 23 export relations. He explains manufactured exports from 1967 in relation to the GNP of the importing country for 1966 by the difference of pci in 1966 and the distance between the trading countries. He did not use a loglinear model. His results were:

$$\frac{X_{ijk}}{Y_i} = a_i + b_i\left(\left|\frac{Y_j}{N_j} - \frac{Y_i}{N_i}\right|\right) + c_i\ D_{ij} \qquad i = 1,2, \ \ldots 23$$

| parameter | mean | number of positive estimates insign. | sign. at 10% | number of negative estimates insign. | sign. at 10% |
|---|---|---|---|---|---|
| $b_j$ | -7.47 | 2 | 0 | 14 | 7 |
| $c_j$ | -4.53 | 0 | 0 | 9 | 14 |

Yamazawa used average data from 1960-'62. He considered 15 groups of countries and allowed for trade within the groups. His results were:

$$\log\left(\frac{X_{ij}/X_{i.}}{X_{.j}/X_{..}}\right) = \underset{}{-0.2556} + \underset{(0.2394)}{1.6534}\log C_{ij} - \underset{(0.0684)}{0.5052}\log\left(\frac{D_{ij}}{14\sqrt{D_{.j}}\sqrt{D_{i.}}}\right)$$

$$+ \underset{(0.1296)}{0.0146}P_1 + \underset{(0.0860)}{0.5042}P_2 + \underset{(0.1950)}{0.7974}P_3 + \underset{(0.0687)}{0.2588}P_4$$

$C_{ij}$ indicates whether "countries" i and j are more ($C_{ij}$ above unity) or less complementary ($C_{ij}$ below unity).
$X_{i.} = \sum_j X_{ij}$ and so on.

$P_k$ are dummy variables; $P_1$ is 1 if both i and j belong to the EEC, $P_1$ is also 1 if if both countries are members of the EFTA; $P_2$ is 1 if both countries belong to the British Commonwealth or the French Community; $P_3$ is 1 if both are socialist; $P_4$ is 1 if i is socialist (capitalist) and j is capitalist (socialist).
Sharma estimated 13 export equations. Each estimation was based upon 54 trade flows of manufactured goods. His results tried to establish a relation between the difference in pci and the trade flows. His estimation was based upon 1968 figures. The results were:

$$\log X_{ijk} = a_{1i} + a_{2i} \log\left(\left|\frac{Y_i}{N_i} - \frac{Y_j}{N_j}\right|\right) + a_{3i} \log D_{ij} + a_{4i} \log Y_j$$
$$+ a_{5i} \log \left(\frac{Y_j}{N_j}\right)$$

| parameter | mean | number of positive estimates insign. | sig. at 10 % | number of negative estimates insign. | sign. at 10 % |
|---|---|---|---|---|---|
| $a_1$ | 3.77 | | | | |
| $a_2$ | -0.01 | 4 | 1 | 7 | 1 |
| $a_3$ | -0.66 | 0 | 0 | 2 | 11 |
| $a_4$ | 0.63 | 0 | 13 | 0 | 0 |
| $a_5$ | 0.10 | 5 | 6 | 2 | 0 |

Pelzman estimated trade flows within the CMEA and between CMEA countries and other countries. He used pooled data from 1954-'70. His results were:

$$\log X_{ij} = 6.72 + \underset{(0.03)}{0.788} \log Y_j + \underset{(0.03)}{0.954} \log Y_i - \underset{(0.04)}{0.177} \log N_j$$
$$- \underset{(0.04)}{0.283} \log N_i - \underset{(0.03)}{1.229} \log D_{ij} + \underset{(0.10)}{2.788} \log P_{ij}$$

$P_{ij}$ indicates whether country i and j are CMEA countries ($P_{ij} = 2$) or not ($P_{ij} = 1$).
As a last example we present a study from Aitken who was interested in enlarging trade within the EEC and the trade within the EFTA. To this end he estimated trade equations for the years 1951 up to

1967 in which dummy variables for the intra EEC trade and intra EFTA trade were included. The parameter for these latter coefficients showed a remarkable increase. As an example the equation based upon 1967 data:

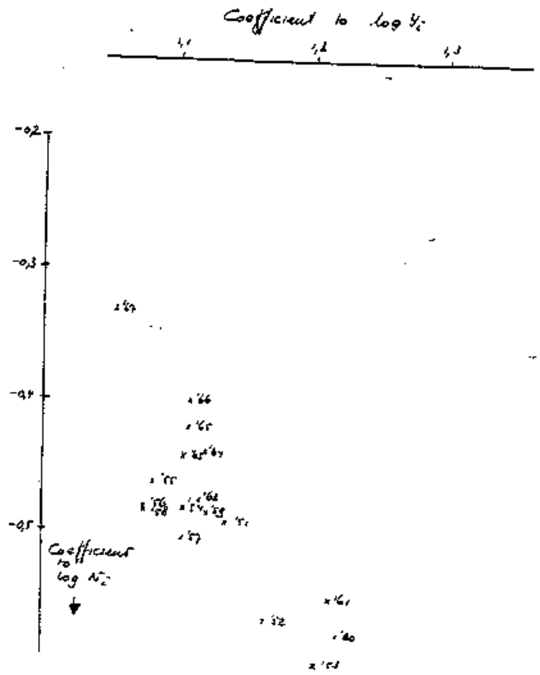
$$\log X_{ij} = \underset{}{1.067} + \underset{(0.10)}{1.052} \log Y_i + \underset{(0.10)}{0.911} \log Y_j - \underset{(0.11)}{0.331} \log N_i$$

$$- \underset{(0.11)}{0.369} \log N_j - \underset{(0.13)}{0.349} \log D_{ij} + \underset{(0.20)}{0.892} \log A_{ij}$$

$$+ \underset{(0.24)}{0.887} \log P_{ij}^{EEC} + \underset{(0.18)}{0.572} \log P_{ij}^{EFTA}$$

$A_{ij}$ = 2 if both i and j are neighbours; $P_{ij}^{EEC}$ is 2 if both countries are members of the EEC and $P_{ij}^{EFTA}$ is 2 if both trading countries are EFTA members.

All authors have used least squares estimators. Most equations resemble each other: they are loglinear equations using as dependent variables the gross national product, distance and population. Let us discuss briefly the inclusion of both population and the gross national product. To begin with, we return to the last example of the gravity equation: the Aitken research. It has already been said that Aitken estimated the equation for all years between 1951 and 1967. Here we plot the estimates of the coefficients to log $Y_i$ and log $N_i$. The plot suggests a relationship between those coefficients. If the coefficient of log $Y_i$ increases, the coefficient to log $N_i$ becomes more negative.



As another example we present two OLS estimates of merchandise trade flows. For 1981 we got two estimates:

$$\log X_{ij} = -3.647 - 0.726 \log D_{ij} + 0.891 \log Y_i + 0.821 \log Y_j - 0.272 \log P_1 - 0.366 \log P_2 + 0.023 \log P_3 - 0.227 \log P_4 + 0.087 \log P_5 - 0.055 \log P_6$$

$R^2 = 0.637$

$$\log X_{ij} = -3.120 - 0.696 \log D_{ij} + 1.054 \log Y_i - 0.316 \log N_i + 0.877 \log Y_j - 0.081 \log N_j - 0.290 \log P_1 - 0.351 \log P_2 + 0.077 \log P_3 - 0.190 \log P_4 + 0.142 \log P_5 + 0.009 \log P_6$$

$R^2 = 0.662$

$P_{123456}$ are dummy variables which have value 1 if a certain country is a trade partner. (1 stands for Peru, 2 for Turkey et cetera). We used 75 countries in the sample, so that the total of trade flows is 5625. Here, multicollinearity between N and Y is clearly demonstrated: inclusion of population (N) as a variable causes a shift in the estimates of Y.
Another remark can be made about the treatment of zero trade flows: trade flows reported by the statistics as zero. A logarithmic transfromation is not possible. Linnemann omitted the data connected with zero trade flows. Other authors limited the number of countries in such a way that trade flows are positive. Yamazawa adjusted low values of the dependent variable. In the following paragraph we show the danger of (perhaps unknowingly) omitting data connected with zero trade flows.
A final remark can be made about the price variable: no author has included an explicit price variable. In a final paragraph we will discuss this matter.

## Biased OLS results

Judge (1980) shows that OLS estimates can be biased when (perhaps) undefined values of the dependent variable and the corresponding independent variables are not included in the estimation procedure. Often, in the gravity equation all variables are transformed in logarithms. When the trade flow is originally zero, no logarithm is defined. In OLS estimations the undefined data of the dependent variable and the corresponding data of the independent variables are excluded.
As an example several data series are generated. First independent data series are set up. The first independent variable ($x_1$) includes 5500 data in which three different values occur: 0.693, 1.733 and 2.340. The second independent variable ($x_2$) includes 5500 data in which also three different values occur: 0.693, 1.386 and 1.792. Of course these values appear manifold in the actual data series.

Then a stochastic term ($u_t$) is generated. This is generated from a a normal ((N(0, 0.3)) distribution. These variables generate the data series of the dependent variable ($y_t$) according to:

$$\begin{cases} y_t = a_t, a_t > 0 \\ y_t \text{ undefined elsewhere} & (2) \\ a_t = -1.160 + 0.400\, x_{1t} + 0.600\, x_{2t} + u_t, \quad t = 1, 2, 3, \ldots \quad 5500 \end{cases}$$

The number of undefined values for $y_t$ can be thought of as being representative for the number of undefined values which appears if trade flows are transformed into logarithms.
In the OLS estimation the dependent variable is excluded from the procedure when the value is undefined. The corresponding independent variables are also excluded in these cases. The results are:

$$y_t = \underset{(0.022)}{-0.571} + \underset{(0.007)}{0.274}\, x_{1t} + \underset{(0.011)}{0.408}\, x_{2t}$$

The standard errors are written between parentheses. The results are clearly biased. Thus it is suspected that OLS estimates of gravity models are biased.

Tobit estimators use as a starting point the likelihood function in which the values of the independent variable are taken into account even when the dependent variable is undefined. The model to be estimated can be written as:

$$\begin{cases} y_t = a_t,\ a_t > 0 \\ y_t \text{ undefined elsewhere} \\ a_t = x_t'\, \beta + u_t,\ u_t \sim IN(0, \sigma^2) \end{cases}$$

The likelihood function is:

$$\mathcal{L}(\beta, \sigma) = \prod_0 (1 - \Phi(x_t'\beta/\sigma)) \prod_1 \frac{1}{\sigma}\varphi((y_t - x_t'\beta)/\sigma) \tag{3}$$

$\Phi(\varphi)$ is the distribution (density) function of the standard normal variable. $\prod_0$ stands for the product of those t for which $y_t$ is undefined. If $y_t$ is undefined, we know $a_t \leq 0$, hence $u_t \leq -x_t'\beta$. From the distribution function of $u_t$ we know: $P(u_t < -x_t'\beta) = \Phi(-x_t'\beta/\sigma)$ $= (1 - \Phi(x_t'\beta/\sigma))$. $\prod_1$ stands for the product of those t for which $y_t = a_t$, $a_t > 0$.

The loglikelihood function is:

$$\log \mathcal{L}(\beta, \sigma) = \sum_0 \log(1 - \Phi(x_t'\beta/\sigma)) - \sum_1 \frac{1}{2\sigma^2}(y_t - x_t'\beta)^2 - n_1 \log \sigma - \frac{n_1}{2}\log(2\pi) \tag{4}$$

$n_1$ is the number of observations for which $y_t$ is defined. So we have a non linear model. We can estimate $(\beta,\sigma)$ by maximizing $\log \mathcal{L}$ with respect to $\beta$ and $\sigma$. This can be done by an iterative procedure. The outline is, let $\binom{\beta}{\sigma}_n$ a $n^{th}$ step, then $\binom{\beta}{\sigma}_{n+1} = \binom{\beta}{\sigma}_n + t_n P_n \gamma_n$ in which $t_n$ is a step length, $\gamma_n$ the gradient of $\log \mathcal{L}$ and $P_n$ a positive definite matrix. For $P_n$ it is possible to take I, the unity matrix,

or $\left[ \dfrac{\partial^2(\log \mathcal{L})}{\partial\binom{\beta}{\sigma}\,\partial\binom{\beta}{\sigma}'} \right]^{-1}$ or another positive definite matrix.

Some try outs illustrated the need to compute $\gamma$ analytically.

In literature two different short cut estimation procedures are presented.

One, derived by Fair, sets the gradient equal to zero. Those first order conditions are used to arrive at an iterative procedure for the estimation of $\binom{\beta}{\sigma}$. In the example illustrated above, we get:

$$y_t = \underset{(0.021)}{-1.174} + \underset{(0.007)}{0.402}\, x_{1t} + \underset{(0.011)}{0.610}\, x_{2t}$$

Another procedure, the Heckman procedure (see Amemiya( 1984) and Judge) starts by rewriting (3) in:

$$\pounds(\beta,\sigma) = \prod_0 (1-\Phi(x_t'\beta/\sigma)) \prod_1 (\Phi(x_t'\beta/\sigma)) \prod_1 \left( \frac{\varphi(\frac{y_t - x_t'\beta}{\sigma})}{\sigma\,\Phi(x_t'\beta/\sigma)} \right) \qquad (5)$$

The first two products describe a probit model for which $\beta/\sigma$ can be estimated. These estimates are substituted in:

$$y_t = x_t'\tilde{\beta} + \tilde{\sigma}\left(\frac{\varphi(x_t'\beta/\sigma)}{\Phi(x_t'\beta/\sigma)}\right) + \varepsilon_t{}^{*} \qquad (6)$$

It can be proven that $E(\varepsilon_t) = 0$, hence (6) can be estimated, using OLS once $(\beta/\sigma)$ is estimated. In our example this procedure yields:

$$y_t = \underset{(0.063)}{-1.064} + \underset{(0.014)}{0.371}\, x_{1t} + \underset{(0.022)}{0.573}\, x_{2t}$$

The Heckman procedure yields worse results. The standard errors are bigger and the differences between estimated coefficients and their true values are greater.

To become more familiar with the Tobit estimator another experiment has been undertaken. To this end 100 estimates have been made of (2).

* Because no restrictions are placed upon $\tilde{\beta}$ and $\tilde{\sigma}$ (such as $\tilde{\beta} = \beta$ or $\sigma = \tilde{\sigma}$) the model in eq (6) is different from the model in (5) since (5) does not allow for different $\sigma$ and $\beta$ in the first and last part in the equation.

To save computer time, the data series have 20 values. Both the procedure of Heckman and the procedure of Fair are used to estimate the coefficients. The results are:

| Heckman | Fair | True Value |
|---|---|---|
| -1.057<br>(1.900) | -1.174<br>(0.132) | -1.160 |
| 0.361<br>(0.095) | 0.407<br>(0.011) | 0.400 |
| 0.596<br>(0.231) | 0.578<br>(0.043) | 0.600 |

The values between parentheses are standard errors computed from the 100 estimates of these coefficients. As an overall result, the results of the Heckman procedure are worse than the results of the Fair procedure. Two procedures (the methods of Heckman and Fair) are thus tested. In the experiments the method of Fair yielded smaller standard errors of the estimates and most estimates were closer to the true value than the method of Heckman. In both methods the estimates of the coefficients were within an acceptable range of the true value. The Tobit estimator is a good solution to the problem which occurs when values of the dependent variable are undefined. This problem is known in the estimation of trade flows when trade flows are modelled in a loglinear gravity model.

OLS and Tobit estimates

In this paragraph three sets of estimates are presented in table 1. Firstly, OLS estimates are presented. Secondly, estimates are presented derived by the method of Fair. Those estimates are derived by maximizing the following loglikelihood function:

$$\log \pounds(\beta,\sigma) = \sum_0 \log( \Phi ( \frac{\frac{1}{2} - x_t'\beta}{\sigma})) - \sum_1 ( \frac{(y_t - x_t'\beta )^2}{2\sigma^2} )$$

$$- n_1 \log \sigma - \frac{n_1}{2} \log(2\pi) \qquad (7)$$

In the first part of eq. (7) the treshold is seen at which flows are reported as zero or non zero. The treshold is set at ½.
Thirdly, the estimates are presented which are derived by using the method described by Heckman. Those estimates are derived by maximizing the likelihood function with respect to $(\beta/\sigma)$:

$$\pounds(\beta/\sigma) = \prod_0 (1 - \Phi(x_t'\beta/\sigma)) \prod_1 (\Phi(x_t'\beta/\sigma)) \qquad (8)^1$$

followed by an OLS estimation of $(\tilde{\beta},\tilde{\sigma})$ in:

$$y_t = x_t'\tilde{\beta} + \tilde{\sigma}(\frac{\varphi(x_t'\beta/\sigma)}{\Phi(x_t'\beta/\sigma)}) + \varepsilon_t \qquad (8)^2$$

Table 1 presents the estimates of $\tilde{\beta}$ and $\tilde{\sigma}$; an appendix presents the estimates of $(\beta/\sigma)$.

Due to a somewhat different approach it did not seem necessary to include the treshhold value ½.

The results are based upon the trade flows between 51 countries. Leaving out the flows from country i to itself, 2550 trade flows are utilized as data.

The estimates are made from year 1970 to 1983.*

---

*Data sources: for trade flows merchandise trade was taken as reported in "Direction of Trade" by the IMF. The data with respect to GNP were taken from the "World Bank Atlas" and the data with respect to distance were seen as distance from the economic centre in a country to an economic centre in the trade partner country. It was assumed that all trade takes place by sea.

Table 1   OLS estimates 1970-'83 for $b_{0123}$ in $\log X_{ij} = b_0 + b_1 \log Y_i + b_2 \log Y_j + b_3 \log D_{ij}$

| Year | $b_0$ | $b_1$ | s.e. | $b_2$ | s.e. | $b_3$ | s.e. | $R^2$ |
|---|---|---|---|---|---|---|---|---|
| 1970 | -2.090 | 0.650 | 0.018 | 0.667 | 0.017 | -0.543 | 0.028 | 0.648 |
| 1971 | -2.172 | 0.666 | 0.018 | 0.679 | 0.018 | -0.556 | 0.029 | 0.648 |
| 1972 | -2.177 | 0.665 | 0.018 | 0.692 | 0.018 | -0.565 | 0.028 | 0.658 |
| 1973 | -2.989 | 0.735 | 0.017 | 0.769 | 0.017 | -0.539 | 0.027 | 0.695 |
| 1974 | -2.914 | 0.730 | 0.018 | 0.755 | 0.018 | -0.526 | 0.029 | 0.659 |
| 1975 | -3.318 | 0.806 | 0.018 | 0.785 | 0.018 | -0.567 | 0.030 | 0.690 |
| 1976 | -3.218 | 0.789 | 0.019 | 0.802 | 0.019 | -0.606 | 0.033 | 0.657 |
| 1977 | -3.472 | 0.814 | 0.018 | 0.808 | 0.018 | -0.582 | 0.032 | 0.685 |
| 1978 | -3.850 | 0.869 | 0.018 | 0.815 | 0.018 | -0.576 | 0.032 | 0.695 |
| 1979 | -3.925 | 0.853 | 0.018 | 0.852 | 0.018 | -0.589 | 0.032 | 0.702 |
| 1980 | -4.147 | 0.884 | 0.019 | 0.853 | 0.019 | -0.573 | 0.032 | 0.684 |
| 1981 | -4.178 | 0.922 | 0.019 | 0.825 | 0.019 | -0.601 | 0.032 | 0.694 |
| 1982 | -3.997 | 0.931 | 0.018 | 0.808 | 0.018 | -0.642 | 0.031 | 0.709 |
| 1983 | -3.983 | 0.938 | 0.019 | 0.822 | 0.019 | -0.675 | 0.032 | 0.697 |

Tobit estimates - method Fair

| Year | $b_0$ | $b_1$ | s.e. | $b_2$ | s.e. | $b_3$ | s.e. | $\log(\mathcal{L}(\beta,\sigma))$ |
|---|---|---|---|---|---|---|---|---|
| 1970 | -3.288 | 0.999 | 0.022 | 0.971 | 0.022 | -0.678 | 0.037 | -4022 |
| 1971 | -3.529 | 1.026 | 0.022 | 0.994 | 0.022 | -0.659 | 0.037 | -4389 |
| 1972 | -3.570 | 1.010 | 0.022 | 1.016 | 0.022 | -0.653 | 0.037 | -4542 |
| 1973 | -4.351 | 1.097 | 0.022 | 1.109 | 0.022 | -0.638 | 0.037 | -5792 |
| 1974 | -4.396 | 1.117 | 0.024 | 1.097 | 0.024 | -0.638 | 0.040 | -5554 |
| 1975 | -4.692 | 1.198 | 0.025 | 1.104 | 0.025 | -0671 | 0.042 | -6176 |
| 1976 | -4.856 | 1.195 | 0.025 | 1.149 | 0.028 | -0.685 | 0.045 | -6369 |
| 1977 | -4.920 | 1.211 | 0.025 | 1.127 | 0.025 | -0.672 | 0.045 | -6537 |
| 1978 | -5.158 | 1.251 | 0.025 | 1.131 | 0.025 | -0.675 | 0.045 | -6883 |
| 1979 | -5.233 | 1.244 | 0.025 | 1.148 | 0.025 | -0.682 | 0.046 | -7059 |
| 1980 | -5.382 | 1.257 | 0.026 | 1.153 | 0.026 | -0.660 | 0.046 | -7413 |
| 1981 | -5.506 | 1.295 | 0.026 | 1.142 | 0.026 | -0.685 | 0.047 | -7574 |
| 1982 | -5.389 | 1.296 | 0.026 | 1.132 | 0.026 | -0.719 | 0.046 | -7371 |
| 1983 | -3.818 | 1.132 | 0.032 | 1.009 | 0.032 | -0.883 | 0.056 | -5438 |

Table 1 cont'd.

Tobit estimates - method Heckman* - these estimates can be thought of as estimates of $\tilde{\beta}$ and $\tilde{\sigma}^2$ in eq. (6) and (8).

| Year | $b_0$ | $b_1$ | s.e. | $b_2$ | s.e. | $b_3$ | s.e. |
|---|---|---|---|---|---|---|---|
| 1970 | -3.925 | 0.901 | 0.025 | 0.894 | 0.025 | -0.643 | 0.027 |
| 1971 | -3.790 | 0.879 | 0.026 | 0.872 | 0.025 | -0.625 | 0.028 |
| 1972 | -3.785 | 0.867 | 0.024 | 0.886 | 0.024 | -0.627 | 0.027 |
| 1973 | -4.197 | 0.883 | 0.023 | 0.909 | 0.023 | -0.584 | 0.028 |
| 1974 | -4.092 | 0.875 | 0.025 | 0.887 | 0.026 | -0.571 | 0.029 |
| 1975 | -4.095 | 0.904 | 0.026 | 0.868 | 0.024 | -0.596 | 0.030 |
| 1976 | -4.254 | 0.913 | 0.028 | 0.911 | 0.027 | -0.635 | 0.031 |
| 1977 | -4.222 | 0.905 | 0.027 | 0.884 | 0.025 | -0.605 | 0.031 |
| 1978 | -4.418 | 0.937 | 0.028 | 0.873 | 0.025 | -0.595 | 0.031 |
| 1979 | -4.425 | 0.913 | 0.027 | 0.901 | 0.025 | -0.604 | 0.032 |
| 1980 | -4.971 | 0.980 | 0.029 | 0.933 | 0.027 | -0.598 | 0.033 |
| 1981 | -4.586 | 0.969 | 0.028 | 0.865 | 0.026 | -0.612 | 0.032 |
| 1982 | -4.080 | 0.941 | 0.027 | 0.817 | 0.026 | -0.644 | 0.030 |
| 1983 | -4.002 | 0.941 | 0.027 | 0.824 | 0.027 | -0.675 | 0.030 |

---

The method described by Fair yielded the best results. As an example let us turn to a survey which confronts the trade values with simulated results for the year 1978. In table 2 we see the estimates are clearly biased if an OLS estimation procedure is used. The best fit is derived if the Fair procedure is used.

Table 2 confronts actual trade figures for 1978 with simulated trade ( $10^{b_0} Y_i^{b_1} Y_j^{b_2} D_{ij}^{b_3}$ ) where $b_k$ are estimates obtained by OLS, Tobit (Fair) and Tobit (Heckman) procedures. Each table consists of several rows ($T_{i.}$) in which the first element ($T_{i1}$) indicates the number of actual trade flows which belong to a certain class. Eg. $T_{21}$ = 141 indicates that 141 trade flows are reported as amounting to \$ 1 million. The other elements in the row ($T_{ij,\ j>1}$) indicate the amount of simulated trade flows - assuming that the original trade flow is included in the first element of the row. Eg. $T_{22}$ = 1 indicates that one simulation yields the simulation being \$1 million, while the actual value was \$ 1 million.

*) In an appendix following "Data Sources" the probit estimates are given. They can be thought of as estimates for $(\beta/\sigma)$ in the first part in eq. (5) and as estimates of $(\beta/\sigma)$ in eq. $(8)^1$.

TABLE 2 Confrontation of actual values and their simulations (OLS)

| class (actual values) | (amount million $) | class(simulations) 1<br>0 | 2<br>1 | 3<br>2 | 4<br>3-4 | 5<br>5-8 | 6<br>9-16 | 7<br>17-32 | 8<br>33-64 | 9<br>65-128 | 10<br>129-256 | 11<br>257-512 | 12<br>513-1024 | 13<br>1025-2048 | 14<br>2049-4096 | 15<br>4097-8192 | 16<br>8193-16384 | 17<br>16385-32768 | 18<br>> 32768 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 762 | 20 | 161 | 111 | 155 | 113 | 96 | 69 | 34 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 141 | 0 | 13 | 8 | 35 | 28 | 37 | 10 | 7 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 77 | 0 | 1 | 2 | 15 | 30 | 12 | 15 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 102 | 0 | 3 | 7 | 20 | 28 | 22 | 14 | 5 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 148 | 0 | 4 | 9 | 15 | 38 | 44 | 24 | 11 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 142 | 0 | 0 | 9 | 13 | 20 | 35 | 36 | 22 | 5 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 176 | 0 | 1 | 3 | 6 | 15 | 40 | 50 | 48 | 8 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 216 | 0 | 2 | 2 | 4 | 6 | 25 | 48 | 65 | 43 | 17 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 189 | 0 | 0 | 0 | 0 | 4 | 10 | 29 | 48 | 56 | 29 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 168 | 0 | 0 | 0 | 2 | 4 | 4 | 21 | 27 | 48 | 40 | 14 | 7 | 1 | 0 | 0 | 0 | 0 | 0 |
| 11 | 149 | 0 | 0 | 0 | 0 | 2 | 2 | 6 | 17 | 23 | 26 | 43 | 25 | 5 | 0 | 0 | 0 | 0 | 0 |
| 12 | 108 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 4 | 13 | 12 | 33 | 26 | 15 | 3 | 0 | 0 | 0 | 0 |
| 13 | 84 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 5 | 9 | 23 | 14 | 14 | 16 | 0 | 0 | 0 | 0 |
| 14 | 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 6 | 7 | 13 | 12 | 4 | 0 | 0 | 0 |
| 15 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 7 | 6 | 6 | 0 | 0 |
| 16 | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 6 | 5 | 3 | 0 |
| 17 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 18 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| | 2550 | 20 | 185 | 151 | 265 | 288 | 328 | 325 | 291 | 213 | 144 | 136 | 82 | 51 | 38 | 16 | 11 | 6 | 0 |

Table 2 Cont'd Estimations derived by Tobit(Fair)

class(simulations)

| class (actual values) | (amount million $) | 1<br>0 | 2<br>1 | 3<br>2 | 4<br>3-4 | 5<br>5-8 | 6<br>9-16 | 7<br>17-32 | 8<br>33-64 | 9<br>65-128 | 10<br>129-256 | 11<br>257-512 | 12<br>513-1024 | 13<br>1025-2048 | 14<br>2049-4096 | 15<br>4097-8192 | 16<br>8193-16384 | 17<br>16385-32768 | 18<br>> 32768 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 762 | 360 | 171 | 51 | 66 | 50 | 43 | 16 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 141 | 34 | 40 | 25 | 19 | 8 | 8 | 4 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 77 | 8 | 30 | 13 | 8 | 10 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 102 | 14 | 36 | 12 | 17 | 12 | 5 | 4 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 148 | 16 | 28 | 34 | 25 | 24 | 14 | 6 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 142 | 14 | 21 | 16 | 23 | 26 | 20 | 15 | 3 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 176 | 6 | 13 | 14 | 26 | 30 | 38 | 34 | 9 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 216 | 5 | 6 | 9 | 14 | 24 | 47 | 44 | 31 | 22 | 12 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 189 | 0 | 2 | 4 | 6 | 19 | 21 | 32 | 45 | 31 | 18 | 9 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 168 | 1 | 4 | 1 | 4 | 10 | 14 | 23 | 28 | 37 | 28 | 9 | 7 | 2 | 0 | 0 | 0 | 0 | 0 |
| 11 | 149 | 0 | 1 | 2 | 1 | 3 | 5 | 15 | 15 | 14 | 25 | 34 | 20 | 10 | 3 | 1 | 0 | 0 | 0 |
| 12 | 108 | 0 | 0 | 1 | 0 | 1 | 2 | 2 | 8 | 12 | 8 | 28 | 21 | 9 | 9 | 7 | 0 | 0 | 0 |
| 13 | 84 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 2 | 4 | 12 | 15 | 13 | 10 | 9 | 12 | 4 | 0 | 0 |
| 14 | 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | 5 | 4 | 6 | 11 | 9 | 5 | 1 | 0 |
| 15 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 6 | 2 | 6 | 3 | 5 |
| 16 | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 4 | 3 | 7 |
| 17 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 18 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| | 2550 | 458 | 352 | 182 | 209 | 219 | 224 | 197 | 149 | 130 | 111 | 102 | 69 | 38 | 38 | 31 | 19 | 7 | 15 |

Table 2 Cont'd Estimations derived by Tobit (Heckman)

class(simulations)

| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (amount million US $) class (actual values) | | 0 | 1 | 2 | 3-4 | 5-8 | 9-16 | 17-32 | 33-64 | 65-128 | 129 - 256 | 257 - 512 | 513 - 1024 | 1025 - 2048 | 2049 - 4096 | 4097 - 8192 | 8193 - 16384 | 16386 - 32768 | > 32768 |
| 1 | 762 | 68 | 187 | 128 | 125 | 93 | 78 | 62 | 18 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 141 | 1 | 17 | 19 | 33 | 33 | 20 | 11 | 5 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 77 | 0 | 3 | 6 | 22 | 22 | 12 | 10 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 102 | 0 | 6 | 11 | 27 | 20 | 23 | 9 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 148 | 0 | 10 | 6 | 22 | 44 | 41 | 17 | 7 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 142 | 0 | 7 | 8 | 13 | 29 | 33 | 27 | 19 | 4 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 176 | 0 | 3 | 3 | 10 | 22 | 40 | 47 | 40 | 7 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 216 | 0 | 3 | 2 | 6 | 10 | 26 | 56 | 59 | 36 | 15 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 189 | 0 | 0 | 0 | 1 | 6 | 17 | 27 | 47 | 53 | 26 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 168 | 0 | 0 | 2 | 1 | 3 | 6 | 24 | 27 | 48 | 37 | 11 | 8 | 1 | 0 | 0 | 0 | 0 | 0 |
| 11 | 149 | 0 | 0 | 0 | 1 | 2 | 2 | 7 | 19 | 19 | 29 | 39 | 26 | 4 | 1 | 0 | 0 | 0 | 0 |
| 12 | 108 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 7 | 11 | 12 | 32 | 25 | 12 | 7 | 0 | 0 | 0 | 0 |
| 13 | 84 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 5 | 9 | 20 | 16 | 13 | 16 | 2 | 0 | 0 | 0 |
| 14 | 46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 5 | 7 | 10 | 14 | 4 | 1 | 0 | 0 |
| 15 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | 6 | 7 | 7 | 0 | 0 |
| 16 | 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 4 | 5 | 5 | 0 |
| 17 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 18 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| | 2550 | 69 | 236 | 185 | 261 | 285 | 299 | 299 | 255 | 191 | 139 | 122 | 84 | 43 | 44 | 17 | 13 | 8 | 0 |

Some more estimates were made but are not reported here. One estimation used a GLS estimator, allowing for an increase in the variance in trade flows if trade takes place between major countries. Those estimates were (like the OLS estimates) clearly biased. Therefore the GLS estimator did not lead to a major improvement in the estimates.
Another estimate was made, using a computer routine which maximized eq. (7). Here the short cuts described by Fair were not utilized. This method was far more expensive in computer time, convincing a researcher with a limited amount of computer time of the benefits of the method of Fair. Fair himself also reported this result.

In literature the most frequently used estimator for the gravity equation is the OLS estimator. This may be due to the fact that Tobit estimators consume much computer time even when a short cut method is utilized such as the Fair method. One example of the Tobit estimator can be found in a thesis, written by Bikker (1982). His likelihood function differed somewhat because he separated two processes: one decision making process ( $d_t = x_t'\xi + v_t$, $v_t \sim IN(0,1)$) whether to trade ( $d_t > 0$ ) or not ( $d_t \leq 0$ ) to trade and a process -given the decision to trade- to decide the amount of trade:($y_t$ amount of trade)

$$
\left|
\begin{array}{l}
y_t = x_t'\beta + u_t \\
v_t = (^{\rho}/_{\sigma})u_t + \varepsilon_t \\
u_t \sim IN(0, \sigma^2) \\
\varepsilon_t \sim IN(0, 1-\rho^2)
\end{array}
\right.
$$

Both processes are based upon the same set of variables ($x_t$)and both **processes** are related (using the correlationcoefficient $\rho$ ).
His loglikelihood function is:

$$\log \pounds(\xi,\rho,\beta,\sigma) = \sum_0 \log \Phi(-x_t'\xi) + \sum_1 \log\left(\frac{x_t'\xi + \frac{\rho}{\sigma}(y_t - x_t'\beta)}{\sqrt{(1-\rho^2)}}\right) - \frac{1}{2\sigma^2}\sum_1 (y_t - x_t'\beta)^2 - n_1 \log\sigma - \frac{n_1}{2}\log(2\pi) \qquad (9)$$

An application of the estimates of the coefficients

The 51 countries for which the trade flows were estimated were not engaged in debt negotiations during the period 1970 - '83. If debt negotiations are used as a proxy for a situation in which severe balance of payments (BOP) problems exist, then it is feasible to use the estimates of trade flows to obtain an estimate for the amount of trade which would have taken place if no BOP problem would have occured. To correct for the particularities of an economy under investigation an array of estimates is made for the years 1970 up to 1983 to give an insight into the development of the estimates. 74 flows between an economy and their trade partners are thus simulated. The simulations are confronted with 74 actual trade flows, from which a ratio is calculated:

$$\text{ratio} = \frac{X_{\cdot j}}{\sum_{\ell} 10^{b_0} Y_{\ell}^{b_1} Y_j^{b_2} D_{\ell j}^{b_3}} \qquad (10)$$

in which $b_{0123}$ are estimates obtained for the trade flow between country i and j, derived by the Tobit estimator using the method of Fair. For example, in case of Brazil import ratios were calculated. See table 3[a].

Table 3[a] Import ratios for Brazil (actual GNP used)

| Year | ratio |
|---|---|
| 1970 | 0.83 |
| 1971 | 0.83 |
| 1972 | 0.87 |
| 1973 | 0.71 |
| 1974 | 1.01 |
| 1975 | 0.79 |
| 1976 | 0.58 |
| 1977 | 0.46 |
| 1978 | 0.47 |
| 1979 | 0.51 |
| 1980 | 0.45 |
| 1981 | 0.46 |
| 1982 | 0.41 |
| 1983 | 0.63 |

The array of ratios shows a steady decrease from the mid seventies towards 1982 in which year the BOP constraint was severly felt in Brazil. But the ratio shows a strange upturn in 1982/83. It seems an improvement in imports.

A closer look reveals that Brazils GNP decreased in 1982/83. Therefore the ratio increases, a statement which can easily been verified by looking at eq. 10.

If an estimate of the trade flow is desired as if no constraint had occurred, the GNP of the country under investigation must be estimated as if no BOP problem had occurred.

A first attempt is to regress the GNP upon GNP of comparable countries during the period in which the country did not experience BOP problems. Then a forecast is made for the period in which the BOP problem occurred. The latter forecast is then used in eq. (10). This manoeuvre yielded better results. This can be seen in table 3[b] in which the strange upturn does not occur in 1982/83.

Table 3[b] Import ratios for Brazil
(forecasts for GNP used)

| Year | ratio |
|---|---|
| 1970 | 0.81 |
| 1971 | 0.78 |
| 1972 | 0.85 |
| 1973 | 0.75 |
| 1974 | 1.10 |
| 1975 | 0.87 |
| 1976 | 0.80 |
| 1977 | 0.60 |
| 1978 | 0.54 |
| 1979 | 0.57 |
| 1980 | 0.56 |
| 1981 | 0.49 |
| 1982 | 0.40 |
| 1983 | 0.39 |

Although some good results were obtained by the above method, the method is not satisfactory. In future research a more elaborate method must be used in order to get a better description of what takes place within a country during BOP constraints and why the GNP declines.

## Inclusion of prices

In the paragraph dealing with the review of literature it becomes apparent that no author used a price variable in the gravity equation. Linnemann (p. 44) showed in his thesis that this procedure presupposes a constraint upon the price elasticity.

His model from which the gravity equation was derived as a reduced form equation will be briefly discussed here. His structural equations were:

$$X_{ij}^{D} = \gamma\, Y_j^{\delta}\, p_i^{\varepsilon}\, D_{ij}^{\rho} \qquad (i \neq j) \qquad (11)$$

$$X_i^{SF} = \omega\, Y_i^{\sigma}\, p_i^{\pi} \qquad (12)$$

$$X_i^{SF} = \sum_{\substack{j \\ j\neq i}} X_{ij}^{D} \qquad (13)$$

$X_{ij}^{D}$ is the demand for merchandise trade and $X_i^{SF}$ is the supply of merchandise trade.

So the price is:

$$p_i = \left( \frac{\gamma}{\omega} \; Y_i^{-\sigma} \sum_{\substack{j \\ j \neq i}} Y_j^{\delta} D_{ij}^{\rho} \right)^{\frac{1}{\pi-\varepsilon}} \tag{14}$$

and the reduced form is:

$$X_{ij} = \gamma^{\frac{\pi}{\pi-\varepsilon}} \; \omega^{\frac{-\varepsilon}{\pi-\varepsilon}} \; Y_i^{\frac{-\sigma\varepsilon}{\pi-\varepsilon}} \; Y_j^{\delta} \; D_{ij}^{\rho} \left( \sum_{\substack{j \\ j \neq i}} Y_j^{\delta} D_{ij}^{\rho} \right)^{\frac{\varepsilon}{\pi-\varepsilon}} \tag{15}$$

Only assuming that $\left(\sum_{j \neq i} Y_j^{\delta} D_{ij}^{\rho}\right)^{\frac{\varepsilon}{\pi-\varepsilon}} = 1$, eq. (15) has the same form as the gravity equation.

Dropping the latter assumption, an iterative procedure yields estimates for the parameters. Each step includes a Tobit estimation ( Fair). The $n^{th}$ iteration is an estimation of $X_{ij}$ by substituting

$p_i = \left(\sum_{\substack{j \\ j \neq i}} Y_j^{\delta_{n-1}} D_{ij}^{\rho_{n-1}}\right)^{\left(\frac{\varepsilon}{\pi-\varepsilon}\right)_{n-1}}$ in eq. (15) in order to estimate

$\left(\gamma^{\frac{\pi}{\pi-\varepsilon}} \omega^{\frac{-\varepsilon}{\pi-\varepsilon}}\right)_n, \frac{\varepsilon\sigma}{\pi-\varepsilon}_n, \delta_n, \rho_n, \left(\frac{\varepsilon}{\pi-\varepsilon}\right)_n$

The estimates are presented in table 4.

<u>Table 4  Estimates with a price variable</u>

| Year | constant | $Y_i$ | s.e. | $Y_j$ | s.e. | $D_{ij}$ | s.e. | $P_i$ | s.e. |
|---|---|---|---|---|---|---|---|---|---|
| 1970 | -3.095 | 1.008 | 0.024 | 0.971 | 0.022 | -0.685 | 0.038 | -0.071 | 0.090 |
| 1971 | -3.264 | 1.037 | * | 0.994 | * | -0.668 | * | -0.092 | * |
| 1972 | -3.053 | 1.031 | 0.025 | 1.015 | 0.022 | -0.669 | 0.038 | -0.173 | 0.093 |
| 1973 | -4.305 | 1.099 | * | 1.109 | * | -0.639 | * | -0.013 | * |
| 1974 | -4.163 | 1.125 | 0.028 | 1.096 | 0.024 | -0.644 | 0.042 | -0.067 | 0.103 |
| 1975 | -4.816 | 1.194 | * | 1.104 | * | -0.668 | * | 0.035 | * |
| 1976 | -5.313 | 1.181 | 0.028 | 1.150 | 0.025 | -0.673 | 0.046 | 0.121 | 0.106 |
| 1977 | -5.551 | 1.193 | * | 1.128 | * | -0.656 | * | 0.166 | * |
| 1978 | -5.874 | 1.230 | 0.028 | 1.132 | 0.025 | -0.657 | 0.047 | 0.185 | 0.110 |
| 1979 | -6.131 | 1.219 | * | 1.150 | * | -0.658 | * | 0.225 | * |
| 1980 | -6.618 | 1.227 | 0.029 | 1.155 | 0.026 | -0.631 | 0.048 | 0.298 | 0.118 |
| 1981 | -7.052 | 1.256 | 0.029 | 1.144 | 0.026 | -0.646 | 0.048 | 0.379 | 0.115 |
| 1982 | -7.557 | 1.242 | 0.029 | 1.136 | 0.026 | -0.665 | 0.048 | 0.541 | 0.114 |
| 1983 | -7.617 | 1.220 | 0.028 | 1.143 | 0.026 | -0.684 | 0.047 | 0.589 | 0.112 |

*No standard errors calculated

The estimates were computed using the method of Fair. Comparing the results with estimates obtained in table 1 when the price variable was excluded, we see a strong indication for multicollinearity between the constant and the prices.

## Some final remarks

The purpose of this study was to do research on the gravity equation. The main finding was that the Tobit estimator yielded better results for the coefficients in the gravity equation than the OLS estimator. This study was undertaken as a part of a larger project, which tries to indicate the influence from BOP problems upon trade flows. As briefly indicated in "An Application of the estimates of coefficients", this is only possible when a better insight in the influence from the BOP problems upon a country's GNP is obtained.

Literature

Aitken, N.D.,"The effect of the EEC and EFTA on European trade: a temporal cross-section analysis", American Economic Review, 1973, pp 881-892.

Alonso, W., "A theory of movements", in Hansen, N.M.,(ed) "Human Settlement Systems", Cambridge, 1978, pp 197-212.

Amamiya, T., "Tobit models: a survey", Journal of Econometrics, 1984, pp 3-61.

Bennett, R.J., and Haining, R.P.,"Spatial Structure and Spatial Interaction: Modelling Approaches to the Statistical Analysis of Geographical Data", J. Royal Statistical Society, 1985, pp 1-36.

Bikker, J. A., "Vraag-aanbodmodellen voor stelsels van geografisch gespreide markten toegepast op internationale handel en ziekenhuisopnamen in Noord-Nederland",(in Dutch with a summary in English), Amsterdam, 1982.

Fair, R.C., "A note on the computation of the tobit estimator", Econometrica, 1977, pp 1723-1727.

Fortune, J.N., "Some determinants of trade in finished manufactures", Swed. J. of Economics, 1971, pp 311-317.

Hirsh, Z. and Lev, B., "Trade and per capita income differentials: a test of the Burenstam-Linder hypothesis", World Development, 1973, pp 11-17.

Judge, G.G., Griffith, W.E., Hill, R.C., Lee, T., "The theory and practice of econometrics", New York, 1980.

Kemming, H., "Raumwirtschaftstheoretische Gravitationsmodelle - eine Untersuchung ihrer analytischen Grundlagen",(in German), Berlin, 1980.

Linnemann, H., "An econometric Study of International Trade Flows", Amsterdam, 1966.

Nijkamp, P., "Gravity and entropy models: the state of the art", Research Memorandum Vrije Universiteit, 1977.

Pelzman, J., "Trade creation and trade diversion in the council of Mutual Economic Assistance: 1954-70", American Economic Review, 1977, pp 713-722.

Sharma, B.J., "Preference similarity hypothesis: survey and further emperical evidence", Indian Economic Journal, 1984, pp 79-96.

Yamazawa, I., "Structural changes in world trade flows", Hitotsubashi J. of Economics, 1971, pp 11-22.

## Data Sources

International Monetary Fund, "Direction of Trade", Annuals.
--- "International Financial Statistics",Annuals.

Luensee,"Entfernungstabellen, Die Entfernung zwischen den Haubthäfen der Erde", Hamburg, 1958-60.

United States Naval Oceanographic Office, "Distances between Ports", Washington, 1965.

Worldbank, "World Bank Atlas", Washington.

## Appendix

Estimates of the probit model of international trade. These estimates can be thought of as estimates of $(\beta/\sigma)$ in eq. (5). The estimates are utilized in the Heckman procedure in order to estimate the bias in a OLS estimation. The bias is estimated by $(\varphi(x_t'\beta/\sigma)/\Phi(x_t'\beta/\sigma))$. See also eq. $(8)^1$ and $(8)^2$.

| Year | constant | $Y_i$ | $Y_j$ | $D_{ij}$ |
|---|---|---|---|---|
| 1970 | -5.84 | 1.35 | 1.24 | -0.89 |
| 1971 | -6.76 | 1.43 | 1.31 | -0.82 |
| 1972 | -6.84 | 1.37 | 1.31 | -0.75 |
| 1973 | -7.85 | 1.44 | 1.40 | -0.70 |
| 1974 | -6.96 | 1.33 | 1.21 | -0.71 |
| 1975 | -7.38 | 1.40 | 1.18 | -0.69 |
| 1976 | -8.38 | 1.49 | 1.31 | -0.69 |
| 1977 | -7.70 | 1.42 | 1.18 | -0.68 |
| 1978 | -7.51 | 1.37 | 1.15 | -0.69 |
| 1979 | -7.45 | 1.36 | 1.08 | -0.65 |
| 1980 | -7.64 | 1.33 | 1.10 | -0.61 |
| 1981 | -7.80 | 1.33 | 1.12 | -0.62 |
| 1982 | -7.72 | 1.27 | 1.11 | -0.57 |
| 1983 | -7.94 | 1.26 | 1.17 | -0.55 |

Serie Research Memoranda (1)

| | | |
|---|---|---|
| 1981-1 | E. Vogelvang | A quarterly econometric model for the Price Formation of Coffee on the World Market |
| 1981-2 | H.P. Smit | Demand and Supply of Natural Rubber, Part I |
| 1981-3 | R. Vos | The political Economy of the Republic of Korea; A proposal for a model framework of an open economy in the ESCAP-region, with emphasis on the Role of the State |
| 1981-4 | F.C. Palm | Structural Econometric Modeling and Time Series Analysis - Towards an Integrated Approach |
| 1981-5 | P. Nijkamp in co-op. with H.v. Handenhoven and R. Janssen | Urban Impact Analysis in a Spatial Context: Methodology and Case Study |
| 1981-6 | R. Ruben | Primaire exporten en ekonomiese ontwikkeling |
| 1981-7 | D.A. Kodde | Het genereren en evalueren van voorspellingen van omzet en netto winst: een toegepast kwantitatieve benadering |
| 1981-8 | B. Out | Financiële vraagstukken onder onzekerheid |
| 1981-9 | P.van Dijck and H. Verbruggen | A Constant-Market-Shares Analysis of ASEAN Manufactured Exports to the European Community |
| 1981-10 | P. Nijkamp, H.de Graaff and E. Sigar | A Multidimensional Analysis of Regional Infrastructure and Economic Development |
| 1981-11 | P. Nijkamp | International Conflict Analysis |
| 1981-12 | A.J. Mathot | L'Utilisation du Crédit lors de l'Achat d'une Voiture |
| 1981-13 | S.P. van Duin en P.A. Cornelis | Onderzoek naar levensomstandigheden en opvattingen over arbeid bij mensen zonder werk, deel I |
| 1981-14 | W. van Lierop and P. Nijkamp | Disaggregate Models of Choice in a Spatial Context |
| 1981-15 | Hidde P. Smit | The World Vehicle Market |
| 1981-16 | F.C. Palm | Structural Econometric Modeling and Time Series Analysis: An Integrated Approach |
| 1981-17 | F.C. Palm and Th.E. Nijman | Linear Regression Using Both Temporally Aggregated and Temporally Disaggregated Data |
| 1981-18 | F.C. Palm and J.M. Sneek | Some econometric Applications of the exact Distribution of the ratio of Two Quadratic Forms in Normal Variates |